To appear in Journal of Applied Probability 49.4 (December 2012)

On the functional CLT for reversible Markov Chains with nonlinear growth of the variance

Martial Longla, Costel Peligrad and Magda Peligrad[1]

Department of Mathematical Sciences, University of Cincinnati, PO Box 210025, Cincinnati, Oh 45221-0025, USA.

Email: martiala@mail.uc.edu, peligrc@ucmail.uc.edu, peligrm@ucmail.uc.edu

Key words: Maximal inequality; reversible processes; Markov chains; martingale approximation; tightness; functional central limit theorem.

Mathematical Subject Classification (2000): 60 F 17, 60 G 05, 60 G 10.

Abstract

In this paper we study the functional central limit theorem for stationary Markov chains with self-adjoint operator and general state space. We investigate the case when the variance of the partial sum is not asymptotically linear in n, and establish that conditional convergence in distribution of partial sums implies functional CLT. The main tools are maximal inequalities that are further exploited to derive conditions for tightness and convergence to the Brownian motion.

1 Introduction

Kipnis and Varadhan (1986) showed that for an additive functional zero mean S_n of a stationary reversible Markov chain the condition $var(S_n)/n \to \sigma^2$ implies convergence of $S_{[nt]}/\sqrt{n}$ to the Brownian motion (here, $[nt]$ is the integer part of nt). There is a considerable amount of papers that further extend and apply this result to infinite particle systems, random

[1]Supported in part by a Charles Phelps Taft Memorial Fund grant and the NSA grant H98230-11-1-0135.

walks, processes in random media, Metropolis-Hastings algorithms. Among others, Kipnis and Landim (1999) considered interacting particle systems, Tierney (1994) discussed the applications to Markov Chain Monte Carlo. Wu (1999) and Zhao and Woodroofe (2008) studied the law of the iterated logarithm, Derriennic and Lin (2001) and Cuny and Peligrad (2010) investigated the central limit theorem started at a point.

Recently, Zhao et al. (2010) addressed the conditional central limit theorem question under the weaker condition $var(S_n) = nh(n)$, where h is a slowly varying function (i.e. $\lim_{n\to\infty} h(nt)/h(n) = 1$ for all $t > 0$). They showed by example the surprising result that the distribution of $S_{[nt]}/stdev(S_n)$ needs not converge to the standard normal distribution in this case. They developed sufficient conditions for convergence to a (possibly non-standard) normal distribution imposed to an approximating martingale.

In this paper we address the question of functional central limit theorem for the case considered by Zhao et al. (2010). Our goal is to establish sufficient conditions imposed on the original sequence. We also show that for reversible Markov chains conditional convergence in distribution of partial sums properly normalized implies functional CLT. The main tools to prove this result are new maximal inequalities based on a triangular forward-backward martingale decomposition and tightness results.

Our paper is organized as follows: Section 2 contains the definitions, a short background of the problem and the results. Section 3 is devoted to the proofs. Section 4 contains a functional central limit theorem for an additive functional associated to a Metropolis-Hastings algorithm, with the variance of partial sums behaving asymptotically like $nh(n)$ (where h is a slowly varying function). All throughout the paper $\Rightarrow$ denotes weak convergence, $[x]$ is the integer part of x and $\to^{\mathbb{P}}$denotes convergence in probability. The notation $a_n \sim b_n$ means $a_n/b_n \to 1$ as $n \to \infty$; $a_n = o(b_n)$ means $a_n/b_n \to 0$ as $n \to \infty$.

2 Definitions, background and results

We assume that $(\xi_n)_{n\in\mathbb{Z}}$ is a stationary Markov chain defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in a general state space $(S, \mathcal{A})$. The marginal distribution is denoted by $\pi(A) = \mathbb{P}(\xi_0 \in A)$. Assume there is a regular conditional distribution for ξ_1 given ξ_0 denoted by

$Q(x,A)=\mathbb{P}(\xi_1\in A|\,\xi_0=x)$. Let Q also denote the Markov operator acting via $(Qf)(x)=\int_S f(s)Q(x,ds)$. Next, let $\mathbb{L}_0^2(\pi)$ be the set of measurable functions on S such that $\int f^2 d\pi<\infty$ and $\int f d\pi=0$. For some function $f\in\mathbb{L}_0^2(\pi)$, let

$$X_i=f(\xi_i),\ S_n=\sum_{i=1}^n X_i,\ \sigma_n=(\mathbb{E}S_n^2)^{1/2}. \tag{1}$$

Denote by $\mathcal{F}_k$ the σ–field generated by ξ_i with $i\le k$.

For any integrable random variable X we denote $\mathbb{E}_k(X)=\mathbb{E}(X|\mathcal{F}_k)$. Under this notation, $\mathbb{E}_0(X_1)=(Qf)(\xi_0)=\mathbb{E}(X_1|\xi_0)$. We denote by $||X||_p$ the norm in $\mathbb{L}_p(\Omega,\mathcal{F},\mathbb{P})$.

The Markov chain is called reversible if $Q=Q^*$, where Q^* is the adjoint operator of Q. The condition of reversibility is equivalent to requiring that (ξ_0,ξ_1) and (ξ_1,ξ_0) have the same distribution or

$$\int_A Q(\omega,B)\pi(d\omega)=\int_B Q(\omega,A)\pi(d\omega)$$

for all Borel sets $A,B\in\mathcal{A}$.

Kipnis and Varadhan (1986) assumed that

$$\lim_{n\to\infty}\frac{\sigma_n^2}{n}=\sigma_f^2 \tag{2}$$

and proved that for any reversible Markov chain defined by (1) this condition implies

$$\frac{S_{[nt]}}{\sqrt{n}}\Rightarrow|\sigma_f|W(t), \tag{3}$$

where $W(t)$ is the standard Brownian motion.

Recently Zhao et al. (2010) analyzed the case when

$$\sigma_n^2=nh(n),\ \text{with } h \text{ a slowly varying function.} \tag{4}$$

In their Proposition 1, they showed that without loss of generality, one can assume that $h(n)\to\infty$, since otherwise either (2) holds, (and this case is already known) or $2S_n=(1+(-1)^{n-1})X_1$ *a.s.* Then, in their Proposition 2 they showed that the representation (4) implies

$$||\mathbb{E}_0(S_n)||_2=o(\sigma_n). \tag{5}$$

On the other hand it is well known that (5) implies (4); see for instance Lemma 1 in Wu and Woodroofe (2004). Therefore, we can state Proposition 2 in Zhao et al. (2010) as follows:

Proposition 1 *For a stationary reversible Markov chain* $(X_n)_{n\in\mathbb{Z}}$ *defined by (1), the relations (4) and (5) are equivalent.*

In their Corollary 2, Zhao et al. (2010) gave sufficient conditions for the validity of the conditional CLT in terms of conditions imposed on the differences of an approximating martingale. In addition, they provided an example of reversible Markov chain satisfying (4), for which the central limit theorem holds with a different normalization.

Throughout this paper we shall assume that $\sigma_n^2 \to \infty$.

By conditional convergence in distribution, denoted by $Y_n|\mathcal{F}_0 \Rightarrow Y$, we understand that for any function g which is continuous and bounded

$$\mathbb{E}_0(g(Y_n)) \to^{\mathbb{P}} \mathbb{E}g(Y) \text{ as } n \to \infty.$$

In other words, let $\mathbb{P}^x$ be the probability associated with the Markov chain started from x and let $\mathbb{E}^x$ be the corresponding expectation. Then, for any $\varepsilon > 0$

$$\pi\{x : |\mathbb{E}^x g(Y_n) - \mathbb{E}g(Y)| > \varepsilon\} \to 0.$$

One of our results is the following invariance principle for functionals of stationary reversible Markov chains. Define

$$W_n(t) = \frac{S_{[nt]}}{\sigma_n}$$

Theorem 2 *Assume* $(\xi_n)_{n\in\mathbb{Z}}$ *is a stationary reversible Markov chain as defined above. Define* $(X_i)_{i\in\mathbb{Z}}$ *by (1) and assume that (4) is satisfied and* S_n/σ_n *is conditionally convergent in distribution to* L*. Then,*

$$W_n(t) \Rightarrow cW(t), \tag{6}$$

where $W(t)$ *is a standard Brownian motion and* c *is the standard deviation of* L*.*

Theorem 2 does not require special properties of the Markov chain such as irreducibility and aperiodicity. However, if these properties are satisfied we have the following simplification:

Corollary 3 *Assume $(\xi_n)_{n\in\mathbb{Z}}$ is a stationary, reversible, irreducible and aperiodic Markov chain such that (4) is satisfied. Then $S_n/\sigma_n \Rightarrow L$ implies (6).*

The proof of Theorem 2 requires the development of several tools. First, we shall establish maximal inequalities that have interest in themselves. As in the Doob maximal inequalities for martingales, we shall compare moments and tail distributions of the maximum of partial sums with those of the corresponding partial sums.

Proposition 4 *Let $(X_i)_{i\in\mathbb{Z}}$ be defined by (1) and $Q = Q^*$. Let $p > 1$ and $q > 1$ such that $1/p + 1/q = 1$. Then for all $n \geq 1$,*

$$|| \max_{1\leq i\leq n} |S_i| \; ||_p \leq || \max_{1\leq i\leq n} |X_i| \; ||_p + (4q+3) \max_{1\leq i\leq n} ||S_i||_p. \tag{7}$$

Remark 5 *Let $p = 2$. Since $(X_i)_{i\in\mathbb{Z}}$ is stationary it is well known that*

$$|| \max_{1\leq i\leq n} |X_i| \; ||_2 = o(n^{1/2}) \text{ as } n \to \infty.$$

If we assume in addition $\liminf_n \sigma_n^2/n > 0$ we deduce that there exists $C > 0$, such that

$$|| \max_{1\leq i\leq n} |S_i| \; ||_2 \leq C \max_{1\leq i\leq n} ||S_i||_2.$$

For the proof of tightness, it is also convenient to have inequalities for the tail probabilities of partial sums. We shall also establish:

Proposition 6 *Let $(X_i)_{i\in\mathbb{Z}}$ be defined by (1) and $Q = Q^*$. Then, for every $x > 0$ and $n \geq 1$,*

$$\mathbb{P}(\max_{1\leq i\leq n} |S_i| > x) \leq \frac{2}{x}[18\mathbb{E}|S_n|I(|S_n| > x/12) + 55 \max_{1\leq i\leq n} ||\mathbb{E}_0(S_i)||_1 + || \max_{1\leq i\leq n} |X_i| \; ||_1].$$

An important step in the proof of Theorem 2 is the use of tightness conditions. We shall give two necessary conditions for tightness, that will ensure continuity of every limiting process.

Proposition 7 *Assume X_i is defined by (1), condition (4) is satisfied and one of the following two conditions holds:*

1. $(S_n^2/\sigma_n^2)_{n\geq 1}$ *is uniformly integrable;*

2. S_n/σ_n *is convergent in distribution.*

Then, $W_n(t)$ *is tight in* $D(0,1)$ *endowed with uniform topology and any limiting process is continuous.*

Finally, we give sufficient conditions for convergence to the standard Brownian Motion.

Proposition 8 *Assume* $(\xi_n)_{n\in\mathbb{Z}}$ *is a stationary reversible Markov chain. Define* $(X_i)_{i\in\mathbb{Z}}$ *by (1) and assume that (5) is satisfied. Assume* $(S_n^2/\sigma_n^2)_{n\geq 1}$ *is uniformly integrable and*

$$\lim_{n\to\infty} \frac{||\mathbb{E}_0(S_n^2) - \sigma_n^2||_1}{\sigma_n^2} = 0. \tag{8}$$

Then,

$$W_n(t) \Rightarrow W(t).$$

3 Proofs

We start with a preliminary martingale decomposition that combines ideas from Wu and Woodroofe (2004) with forward-backward martingale approximation of Meyer and Zheng (1984) and Lyons and Zheng (1988).

3.1 Forward-backward martingale decomposition

As in Wu and Woodroofe (2004) for $n \geq 1$ fixed, define the stationary sequences

$$\theta_k^n = \frac{1}{n}\sum_{i=0}^{n-1} \mathbb{E}_k(X_k + ... + X_{k+i}), \text{ and}$$

$$D_k^n = \theta_k^n - \mathbb{E}_{k-1}(\theta_k^n). \tag{9}$$

Then, $(D_k^n)_{k\in Z}$ is a triangular array of martingale differences adapted to the filtration $\mathcal{F}_n = \sigma(\xi_i, i \leq n)$. Notice that

$$\begin{aligned}\theta_k^n &= X_k + \frac{1}{n}\sum_{i=1}^{n-1}\mathbb{E}_k(S_{k+i} - S_k) = X_k + \mathbb{E}_k(\theta_{k+1}^n) - \frac{1}{n}\mathbb{E}_k(S_{k+n} - S_k)\\ &= X_k + \theta_{k+1}^n - D_{k+1}^n - \frac{1}{n}\mathbb{E}_k(S_{k+n} - S_k).\end{aligned}$$

Therefore,

$$X_k = D_{k+1}^n + \theta_k^n - \theta_{k+1}^n + \frac{1}{n}\mathbb{E}_k(S_{k+n} - S_k). \tag{10}$$

We construct now a martingale approximation for the reversed process adapted to the filtration $\mathcal{G}_n = \sigma(\xi_i, i \geq n)$. We introduce the notation $\tilde{\mathbb{E}}_1(X_0) = \mathbb{E}(X_0|\mathcal{G}_1) = \mathbb{E}(X_0|\xi_1) = (Q^* f)(\xi_1)$.

Now, let

$$\tilde{\theta}_k^n = \frac{1}{n}\sum_{i=0}^{n-1} \tilde{\mathbb{E}}_k(X_{k-i} + ... + X_k).$$

With this notation

$$X_{k+1} = \tilde{D}_k^n + \tilde{\theta}_{k+1}^n - \tilde{\theta}_k^n + \frac{1}{n}\tilde{\mathbb{E}}_{k+1}(X_{-n+k+1} + ... + X_k). \tag{11}$$

where $\tilde{D}_k^n$ are martingale differences with respect to the filtration $\mathcal{G}_k = \sigma(\xi_i, i \geq k)$, $\tilde{D}_k^n = \tilde{\theta}_k^n - \mathbb{E}_{k+1}\tilde{\theta}_k^n$.

If we assume that $Q = Q^*$, we have $\tilde{\mathbb{E}}_1(X_0) = \mathbb{E}(X_2|\xi_1) = (Qf)(\xi_1)$. Therefore, $\tilde{\theta}_k^n = \theta_k^n$, $\tilde{\theta}_{k+1}^n = \theta_{k+1}^n$ and $\tilde{\mathbb{E}}_{k+1}(X_{-n+k+1} + ... + X_k) = \mathbb{E}_{k+1}(X_{k+2} + ... + X_{k+n+1})$. Adding relations (10) and (11) leads to

$$X_k + X_{k+1} = D_{k+1}^n + \tilde{D}_k^n + \frac{1}{n}\mathbb{E}_k(S_n - S_k) + \frac{1}{n}\mathbb{E}_{k+1}(S_{k+n+1} - S_{k+1}).$$

Summing these relations we obtain the representation

$$\sum_{i=0}^{k-1}(X_i + X_{i+1}) = \sum_{i=1}^{k}[(D_i^n + \tilde{D}_{i-1}^n) + \frac{1}{n}\mathbb{E}_{i-1}(S_{n+i-1} - S_{i-1}) + \frac{1}{n}\mathbb{E}_i(S_{n+i} - S_i)].$$

So,

$$2S_k + (X_0 - X_k) = \sum_{i=1}^{k}(D_i^n + \tilde{D}_{i-1}^n) + \bar{R}_k^n,$$

where

$$\bar{R}_k^n = \frac{1}{n}\sum_{i=1}^{k}[\mathbb{E}_{i-1}(S_{n+i-1} - S_{i-1}) + \mathbb{E}_i(S_{n+i} - S_i)].$$

Therefore, in the reversible case, we get the following backward-forward martingale representation

$$S_k = \frac{1}{2}[(X_k - X_0) + (M_k^n + \tilde{M}_k^n) + \bar{R}_k^n], \tag{12}$$

where $M_k^n = \sum_{i=1}^k D_i^n$ is a forward martingale adapted to the filtration $\mathcal{F}_k$ and $\tilde{M}_k^n = \sum_{i=0}^{k-1} \tilde{D}_i^n$ is a backward martingale adapted to the filtration $\mathcal{G}_k$.

Also, it is convenient to point out a related martingale approximation, which helps us relate the partial sums with a martingale adapted to the same filtration. Notice that

$$\theta_k^n = X_k + \frac{1}{n}\sum_{i=1}^{n-1} \mathbb{E}_k(S_{k+i} - S_k) = X_k + \bar{\theta}_k^n,$$

$$\text{where} \qquad \bar{\theta}_k^n = \frac{1}{n}\sum_{i=1}^{n-1} \mathbb{E}_k(S_{k+i} - S_k).$$

Starting from (10) and using this notation we obtain

$$X_{k+1} = D_{k+1}^n + \bar{\theta}_k^n - \bar{\theta}_{k+1}^n + \frac{1}{n}\mathbb{E}_k(S_{k+n} - S_k).$$

So, summing these relations, denoting as above $M_k^n = \sum_{i=1}^k D_i^n$, we obtain for every stationary sequence, not necessarily reversible, and for any n and m,

$$S_m = M_m^n + R_m^n, \tag{13}$$

$$\text{where} \quad R_m^n = \bar{\theta}_0^n - \bar{\theta}_m^n + \frac{1}{n}\sum_{k=0}^{m-1} \mathbb{E}_k(S_{k+n} - S_k).$$

3.2 Proof of Proposition 4

We start from (12) and take the maximum on both sides. We easily obtain

$$\max_{1\le i\le n} |S_i| \le \frac{1}{2}(|X_0| + \max_{1\le i\le n} |X_i| + \max_{1\le i\le n} |M_i^n + \tilde{M}_i^n| + \max_{1\le i\le n} |\bar{R}_i^n|). \tag{14}$$

Notice that

$$\max_{1\le i\le n} |\bar{R}_i^n| \le \frac{1}{n}\sum_{i=1}^n (|\mathbb{E}_{i-1}(S_{n+i-1} - S_{i-1})| + |\mathbb{E}_i(S_{n+i} - S_i)|),$$

whence, by Minkowski's inequality and stationarity, for any $p \ge 1$

$$|| \max_{1\le i\le n} |\bar{R}_i^n| \; ||_p \le \frac{2}{n}\sum_{i=1}^n ||\mathbb{E}_i(S_{n+i} - S_i)||_p = 2||\mathbb{E}_0(S_n)||_p. \tag{15}$$

Taking into account that $\max_{1\le k\le n} |\tilde{M}_k^n| \le |\tilde{M}_n^n| + \max_{1\le k\le n} |\tilde{M}_n^n - \tilde{M}_k^n|$, we easily deduce

$$|| \max_{1\le k\le n} |M_k^n + \tilde{M}_k^n| \; ||_p \le || \max_{1\le k\le n} |M_k^n| \; ||_p + || \max_{1\le k\le n} |\tilde{M}_n^n - \tilde{M}_k^n| \; ||_p + ||\tilde{M}_n^n||_p,$$

whence, by Doob maximal inequality applied twice, stationarity and reversibility,

$$|| \max_{1\le k\le n} |M_k^n + \tilde{M}_k^n| \; ||_p \le q||M_n^n||_p + (q+1)||\tilde{M}_n^n||_p = (2q+1)||M_n^n||_p,$$

(where q is the conjugate of p).

From (13) we have $M_n^n = S_n - R_n^n$, and from Minkowski's inequality we deduce that

$$||M_n^n||_p \le ||S_n||_p + \frac{2}{n}\sum_{i=0}^{n-1} ||\mathbb{E}_0(S_i)||_p + ||\mathbb{E}_0(S_n)||_p$$

whence,

$$||M_n^n||_p \le ||S_n||_p + 3\max_{1\le i\le n} ||\mathbb{E}_0(S_i)||_p. \tag{16}$$

From (14), (15) and (16) we deduce the following extension of the Doob maximal inequality for reversible processes:

$$|| \max_{1\le i\le n} |S_i| \; ||_p \le \frac{1}{2}(||X_0||_p + || \max_{1\le i\le n} |X_i| \; ||_p + (2q+1)[||S_n||_p + 3\max_{1\le i\le n} ||\mathbb{E}_0(S_i)||_p] + 2||\mathbb{E}_0(S_n)||_p).$$

Taking now into account that $||\mathbb{E}_0(S_i)||_p \le ||S_i||_p$, Proposition 4 is established. $\Diamond$

3.3 Proof of Proposition 6

For the proof of this proposition we shall use the following claim that can be easily obtained by truncation:

Claim 9 *Let X and Y be two positive random variables. Then for all $x \ge 0$*

$$\mathbb{E}XI(Y > x) \le \mathbb{E}XI(X > x/2) + \frac{x}{2}\mathbb{P}(Y > x).$$

For every $x \ge 0$, using (14), we obtain

$$\mathbb{P}(\max_{1\le i\le n} |S_i| > x) \le \mathbb{P}(\max_{1\le i\le n} |M_i^n + \tilde{M}_i^n| > x) + \mathbb{P}(|X_0| + \max_{1\le i\le n} |X_i| + \max_{1\le i\le n} |\bar{R}_i^n| > x). \tag{17}$$

Applying the Markov inequality, then the triangle inequality followed by (15) with $p = 1$, leads to

$$\mathbb{P}(|X_0| + \max_{1\le i\le n} |X_i| + \max_{1\le i\le n} |\bar{R}_i^n| > x) \le \frac{2}{x}(|| \max_{1\le i\le n} |X_i| \; ||_1 + ||\mathbb{E}_0(S_n)||_1). \tag{18}$$

By triangle inequality and reversibility

$$\mathbb{P}(\max_{1\le i\le n}|M_i^n+\tilde{M}_i^n|>x)\le \mathbb{P}(\max_{1\le i\le n}|M_i^n|>x/3)+$$

$$+\mathbb{P}(\max_{1\le i\le n}|\sum_{k=i}^{n}\tilde{D}_i^n|>x/3)+\mathbb{P}(|\tilde{M}_n^n|>x/3)\le 3\mathbb{P}(\max_{1\le i\le n}|M_i^n|>x/3).$$

Then, by Doob maximal inequality and the above Claim applied to $X=|M_n^n|$ and $Y=\max_{1\le i\le n}|M_i^n|$ we obtain

$$\mathbb{P}(\max_{1\le i\le n}|M_i^n|>x/3)\le \frac{3}{x}\mathbb{E}|M_n^n|I(\max_{1\le i\le n}|M_i^n|>x/3)$$
$$\le \frac{3}{x}\mathbb{E}|M_n^n|I(|M_n^n|>x/6)+\frac{1}{2}\mathbb{P}(\max_{1\le i\le n}|M_i^n|>x/3),$$

implying

$$\mathbb{P}(\max_{1\le i\le n}|M_i^n|>x/3)\le \frac{6}{x}\mathbb{E}|M_n^n|I(|M_n^n|>x/6).$$

Now, we express the right-hand side in terms of S_n. By (13) we have $M_n^n=S_n-R_n^n$ and using the fact that for all positive real numbers x,y,a we have $(x+y)I(x+y>a)\le 2xI(x>a/2)+2yI(y>a/2)\le 2xI(x>a/2)+2y$, we obtain

$$\mathbb{E}|M_n^n|I(|M_n^n|>x/6)\le 2\mathbb{E}|S_n|I(|S_n|>x/12)+2||R_n^n||_1$$
$$\le 2\mathbb{E}|S_n|I(|S_n|>x/12)+6\max_{1\le i\le n}||\mathbb{E}_0(S_i)||_1.$$

Therefore,

$$\mathbb{P}(\max_{1\le i\le n}|M_i^n|>x/3)\le \frac{6}{x}[2\mathbb{E}|S_n|I(|S_n|>x/12)+6\max_{1\le i\le n}||\mathbb{E}_0(S_i)||_1]$$

and so

$$\mathbb{P}(\max_{1\le k\le n}|M_k^n+\tilde{M}_k^n|>x)\le \frac{18}{x}[2\mathbb{E}|S_n|I(|S_n|>x/12)+6\max_{1\le i\le n}||\mathbb{E}_0(S_i)||_1]. \tag{19}$$

Thus, (17), (18) and (19) lead to

$$\mathbb{P}(\max_{1\le i\le n}|S_i|>x)\le \frac{2}{x}[18\mathbb{E}|S_n|I(|S_n|>x/12)+55\max_{1\le i\le n}||\mathbb{E}_0(S_i)||_1+||\max_{1\le i\le n}|X_i|\,||_1].$$

◊

3.4 Proof of Proposition 7

We prove first the conclusion of the proposition under the assumption that $(S_n^2/\sigma_n^2)_{n\geq 1}$ is uniformly integrable.

By stationarity and by Theorem 8.3 in Billingsley (1968) formulated for random elements of D (see page 137 in Billingsley, 1968) we have to show that for all $\varepsilon > 0$

$$\lim_{\delta\to 0^+} \limsup_{n\to\infty} \frac{1}{\delta}\mathbb{P}(\max_{1\leq k\leq [n\delta]} |S_k| > \varepsilon\sigma_n) = 0. \tag{20}$$

By Proposition 6,

$$\mathbb{P}(\max_{1\leq k\leq [n\delta]} |S_k| > \varepsilon\sigma_n) \leq \frac{2}{\varepsilon\sigma_n}[18\mathbb{E}|S_{[n\delta]}|I(|S_{[n\delta]}| > \varepsilon\sigma_n/12)+ \tag{21}$$
$$55 \max_{1\leq i\leq [n\delta]} \mathbb{E}|\mathbb{E}_0(S_i)| + \mathbb{E}\max_{1\leq i\leq n} |X_i|].$$

We shall analyze each term from the right-hand side of inequality (21) separately.

By the fact that $\lim_{n\to\infty} \sigma_{[n\delta]}^2/\delta\sigma_n^2 = 1$, taking into account uniform integrability of $(S_n^2/\sigma_n^2)_{n\geq 1}$ leads to

$$\lim_{\delta\to 0^+} \limsup_{n\to\infty} \frac{1}{\delta\sigma_n}\mathbb{E}|S_{[n\delta]}|I(|S_{[n\delta]}| > \frac{\varepsilon\sigma_n}{12}) \leq$$
$$\lim_{\delta\to 0^+} \limsup_{n\to\infty} \frac{24}{\varepsilon\sigma_n^2}\mathbb{E}S_n^2 I(\frac{|S_n|}{\sigma_n} > \frac{\varepsilon}{24\delta^{1/2}}) = 0.$$

By stationarity and the fact that $\liminf_n \sigma_n^2/n > 0$ we have

$$\frac{1}{\sigma_n^2}(\mathbb{E}\max_{1\leq i\leq n} |X_i|)^2 \leq \frac{1}{\sigma_n^2}\mathbb{E}\max_{1\leq i\leq n} |X_i|^2 \to 0 \text{ as } n\to\infty. \tag{22}$$

Then, by condition (5) and Proposition 1,

$$\frac{1}{\sigma_n^2}\max_{1\leq i\leq [n\delta]} (\mathbb{E}|\mathbb{E}_0(S_i)|)^2 \leq \frac{1}{\sigma_n^2}\max_{1\leq i\leq [n\delta]} \mathbb{E}[\mathbb{E}_0(S_i)]^2 \to 0 \text{ as } n\to\infty. \tag{23}$$

Then, combining the last three convergence results with the inequality (21) leads to (20).

To prove the second part of this proposition, assume now that $S_n/\sigma_n \Rightarrow L$. By Theorem 5.3 in Billingsley (1968), we notice that the limit has finite second moment, namely

$$\mathbb{E}L^2 \leq \liminf_{n\to\infty} ||S_n||_2^2/\sigma_n^2 = 1. \tag{24}$$

Furthermore, since $(|S_n|/\sigma_n)_{n\geq 1}$ is uniformly integrable (because $\mathbb{E}S_n^2/\sigma_n^2 = 1$), by (5) and Theorem 5.4 in Billingsley (1968), it follows that

$$\limsup_{n\to\infty} \frac{1}{\sigma_n}\mathbb{E}|S_{[n\delta]}|I(|S_{[n\delta]}| > \frac{\varepsilon\sigma_n}{12}) \leq \frac{1}{\sqrt{\delta}} \lim_{n\to\infty} \frac{1}{\sigma_{[n\delta]}}\mathbb{E}|S_{[n\delta]}|I(\frac{|S_{[n\delta]}|}{\sigma_{[n\delta]}} > \frac{\varepsilon}{24\sqrt{\delta}}) \tag{25}$$
$$= \frac{1}{\sqrt{\delta}}\mathbb{E}|L|I(|L| > \frac{\varepsilon}{24\sqrt{\delta}}).$$

By passing to the limit in relation (21) and using (22), (23), and (25) we obtain,

$$\limsup_{n\to\infty} \frac{1}{\delta}\mathbb{P}(\max_{1\leq k\leq[n\delta]} |S_k| > \varepsilon\sigma_n) \leq \frac{36}{\varepsilon\delta^{1/2}}\mathbb{E}|L|I(|L| > \frac{\varepsilon}{24\sqrt{\delta}}).$$

Then, clearly

$$\limsup_{n\to\infty} \frac{1}{\delta}\mathbb{P}(\max_{1\leq k\leq[n\delta]} |S_k| > \varepsilon\sigma_n) \leq \frac{36\times 24}{\varepsilon}\mathbb{E}L^2I(|L| > \frac{\varepsilon}{24\sqrt{\delta}}).$$

Finally, taking into account (24), the conclusion follows by letting $\delta \to 0^+$. $\Diamond$

3.5 Proof of Theorem 2

Because conditional convergence in distribution implies weak convergence, it follows that $S_n/\sigma_n \Rightarrow L$. Then, by the second part of Proposition 7, $W_n(t)$ is tight in $C(0,1)$ endowed with uniform topology with all possible limits in $C(0,1)$. Now, let us consider a convergent subsequence, say $W_{n'}(t) \Rightarrow X(t)$. Then $X(t)$ is continuous and since S_n/σ_n is conditionally convergent in distribution, $X(t)$ has independent increments (by the next lemma in this subsection applied on subsequences). It is well known [see, for instance, Doob (1953)] that the process $X(t)$ has the representation $X(t) = at + bW(t)$ for some constants a and b, where $W(t)$ is the standard Brownian motion. Without restricting the generality, by symmetry, we can assume $b > 0$. To identify the constants, we use the convergence of moments in the limit theorem, namely Theorem 5.4 in Billingsley (1968). Notice that $(S_n/\sigma_n)_{n\geq 1}$ is uniformly integrable in $\mathbb{L}_1$ since it is bounded in $\mathbb{L}_2$. We use this remark to obtain

$$\mathbb{E}L = \lim_{n\to\infty} \mathbb{E}S_n/\sigma_n = 0 = \lim_{n'\to\infty} \mathbb{E}W_{n'}(1) = \mathbb{E}X(1) = a + b\mathbb{E}W(1) = a,$$

so $a = 0$. Finally, by the same argument it follows that

$$\lim_{n\to\infty} \mathbb{E}|S_n|/\sigma_n = \mathbb{E}|L| = \lim_{n'\to\infty} \mathbb{E}|W_{n'}(1)| = \mathbb{E}|X(1)| = b\mathbb{E}|W(1)| = b\sqrt{2/\pi}.$$

and so $b = \mathbb{E}|L|\sqrt{\pi/2}$. It follows that $X(t) = (\mathbb{E}|L|\sqrt{\pi/2})W(t)$. In particular it follows that L has normal distribution and therefore $\mathbb{E}|L|\sqrt{\pi/2}$ is the standard deviation of L. $\Diamond$

Lemma 10 *Under the assumptions of Theorem 2, if $W_n(t) \Rightarrow X(t)$, then $X(t)$ has independent increments.*

Proof. Without loss of generality, for simplicity we consider only two increments. For any $0 \le s < t \le 1$, we shall show that

$$(W_n(s), W_n(t) - W_n(s)) \Rightarrow (X(s), X(t) - X(s))$$

where $X(s)$ and $X(t) - X(s)$ are independent. By the Cramér-Wold device it is enough to show that for any two real numbers a and b,

$$A = \mathbb{E}\exp[iaW_n(s) + ib(W_n(t) - W_n(s))]$$
$$-\mathbb{E}\exp[iaX(s)]\mathbb{E}\exp[ib(X(t) - X(s))] \to 0.$$

To see this, notice that

$$\mathbb{E}\exp[iaW_n(s) + ib(W_n(t) - W_n(s))]$$
$$= \mathbb{E}\exp[iaW_n(s)]\mathbb{E}_{[ns]}\exp[ib(W_n(t) - W_n(s))].$$

By adding and substracting $\mathbb{E}\exp[iaW_n(s)]\mathbb{E}\exp[ib(X(t) - X(s))]$ to A, we easily obtain

$$|A| \le \mathbb{E}|\mathbb{E}_{[ns]}\exp[ib(W_n(t) - W_n(s)] - \mathbb{E}\exp[ib(X(t) - X(s))]|+$$
$$+|\mathbb{E}\exp[iaW_n(s)] - \mathbb{E}\exp[iaX(s)]| = I + II.$$

Since we assume that $W_n(s) \Rightarrow X(s)$, it follows that $II \to 0$. Furthermore, by (4), $X(s)$ and $s^{1/2}L$ are identically distributed.

To treat the term I, notice that by stationarity and the definition of $W_n(t)$ we have that

$$I = \mathbb{E}|\mathbb{E}_0\exp[ib(S_{[nt]-[ns]}/\sigma_n)] - \mathbb{E}\exp[ib(X(t) - X(s))]|. \tag{26}$$

Because we assume that $\sigma_n \to \infty$ we have

$$\frac{1}{\sigma_n}\mathbb{E}|S_{[nt]-[ns]} - S_{[n(t-s)]}| \to 0, \tag{27}$$

which easily implies that for all b,

$$\mathbb{E}|\mathbb{E}_0 \exp[ib(S_{[nt]-[ns]}/\sigma_n)] - \mathbb{E}_0 \exp[ib(S_{[n(t-s)]}/\sigma_n)]| \to 0. \tag{28}$$

Now, since $S_{[n(t-s)]}/\sigma_n \Rightarrow X(t-s)$ and $S_{[nt]} - S_{[ns]}/\sigma_n \to X(t) - X(s)$, we deduce from (27) and stationarity that $X(t-s)$ and $X(t) - X(s)$ have the same distribution. Furthermore, by (4), we deduce that $S_{[n(t-s)]}/\sigma_n$ is also conditionally convergent in distribution; so, we also have $X(t-s)$ is distributed as $(t-s)^{1/2}L$. Whence, by taking also into account (26) and (28) it follows that

$$\limsup_{n\to\infty} I = \limsup_{n\to\infty} \mathbb{E}|\mathbb{E}_0 \exp[ib(S_{[n(t-s)]}/\sigma_n)] - \mathbb{E}\exp[ib(X(t-s))]| = 0,$$

leading to the conclusion. $\Diamond$

3.6 Proof of Corollary 3

The proof of this corollary follows the lines of Theorem 2 with the exception that we replace Lemma 10 by the following Lemma:

Lemma 11 *Under the assumptions of Corollary 3, if $W_n(t) \Rightarrow X(t)$, then $X(t)$ has independent increments.*

Proof. We mention that by the fact that the Markov chain is stationary, irreducible and aperiodic it follows that it is absolute regular [see Theorems 21.5 and its Corollary 21.7 in volume 2 of Bradley (2007)]. It is well known that an absolutely regular sequence is strong mixing [see the chart on page 186, volume 1, Bradley (2007)]. This means that $\alpha_n \searrow 0$ where

$$\alpha_n = \sup \mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B);$$

here the supremum is taken over all $A \in \sigma(\xi_i, i \leq 0)$ and $B \in \sigma(\xi_i, i \geq n)$. Because we know from the proof of Theorem 2 that the process $X(t)$ is continuous, it is enough to show that for all k and $0 < s_1 < t_1 < s_2 < t_2 < ... < s_k < t_k < 1$ the increments $(X(t_i) - X(s_i))_{1\leq i\leq k}$ are independent. Now, using the definitions of α_n and $W_n(t)$, we get by recurrence

$$|\mathbb{P}(\cap_{i=1}^k (W_n(t_i - s_i) \in A_i)) - \Pi_{i=1}^k \mathbb{P}(W_n(t_i - s_i) \in A_i)| \leq$$
$$\min_{1\leq i\leq k-1} \alpha_{[n(s_{i+1}-t_i)]} \to 0 \text{ as } n \to \infty,$$

for any Borelians $A_1, ..., A_k$. The conclusion follows by passing to the limit with n. $\Diamond$

3.7 Proof of Proposition 8

By Proposition (1) we know that $\sigma_n^2 = nh(n)$ with h a function slowly varying at infinity. Then, by the first part of Proposition 7, $W_n(t)$ is tight in $D(0,1)$. It remains to apply Theorem 19.4 in Billingsley (1968).

4 Application to a Metropolis-Hastings algorithm

In this section we analyze a standardized example of a stationary irreducible and aperiodic Metropolis-Hastings algorithm with uniform marginal distribution. This type of Markov chain is interesting since it can easily be transformed into Markov chains with different marginal distributions. We point out a central limit theorem under a normalization other than the variance of partial sums. Markov chains of this type are often studied in the literature from different points of view, as in Doukhan et al (1994), Rio (2000 and 2009), Merlevède and Peligrad (2010). The idea of considering Metropolis-Hastings algorithm in this context comes from Zhao et al. (2010).

Let $E = [-1, 1]$. We define now the transition probabilities of a Markov chain by

$$Q(x, A) = (1 - |x|)\delta_x(A) + |x|\upsilon(A)\,,$$

where δ_x denotes the Dirac measure and υ on $[-1, 1]$ satisfies

$$\upsilon(dx) = |x|dx. \tag{29}$$

Then, there is a unique invariant measure, the uniform distribution on $[-1, 1]$,

$$\pi(dx) = dx/2,$$

and the stationary Markov chain $(\xi_i)_i$ with values in E and transition probability $Q(x, A)$ is reversible and positively recurrent. Moreover, for any odd function f we have

$$Q^k(f)(\xi_0) = \mathbb{E}(f(\xi_k)|\xi_0) = (1 - |\xi_0|)^k f(\xi_0) \text{ a.s.} \tag{30}$$

For the odd function $f(x) = sign\ x$, define $X_i = sign\ \xi_i$. In this context we shall show:

Example 12 *Let $(X_j)_{j\geq 1}$ defined above. Then $\sigma_n^2/(2n\log n)\to 1$ and*

$$\frac{1}{\sigma_n}\sum_{j=1}^{[nt]} X_j \Rightarrow \frac{1}{2^{1/2}}W(t).$$

where $W(t)$ is the standard Brownian motion.

Proof. For any $m\geq 0$ we have

$$\mathbb{E}(X_0X_m)=\mathbb{E}(f(\xi_0)Q^m(f)(\xi_0))=\int_E (1-|x|)^m\pi(dx)=1/(m+1).$$

Therefore, by simple computations, we obtain

$$\sigma_n^2\sim 2n\log n \text{ as } n\to\infty.$$

Now, to find the limiting distribution of S_n properly normalized, we study the regeneration process. Let

$$T_0=\inf\{i>0:\xi_i\neq\xi_0\}$$

and

$$T_{k+1}=\inf\{i>T_k:\xi_i\neq\xi_{i-1}\},\quad \tau_k=T_{k+1}-T_k.$$

It is well known that $(\xi_{\tau_k},\tau_k)_{k\geq 1}$ are i.i.d. random variables with ξ_{τ_k} having the distribution υ. Furthermore,

$$\mathbb{P}(\tau_1>n|\xi_{\tau_1}=x)=(1-|x|)^n.$$

Then, it follows that

$$\mathbb{E}(\tau_1|\xi_{\tau_1}=x)=\frac{1}{|x|}\quad\text{and}\quad\mathbb{E}(\tau_1)=2.$$

So, by the law of large numbers $T_n/n\to 2$ a.s.

Let us study the tail distribution of τ_1. Since

$$\mathbb{P}(\tau_1|X_{\tau_1}|>y|\xi_{\tau_1}=x)=\mathbb{P}(\tau_1>y|\xi_{\tau_1}=x)=(1-|x|)^y,$$

by integration we obtain

$$\mathbb{P}(\tau_1>y)=\int_{-1}^1(1-|x|)^y|x|dx=2\int_0^1(1-x)^y x dx\sim 2y^{-2} \text{ as } y\to\infty.$$

Moreover, $\mathbb{E}(\tau_k X_{\tau_k}) = 0$ by symmetry. Also

$$H(y) = \mathbb{E}(\tau_1^2 I(\tau_1 \le y)) \sim 4 \ln y.$$

Define a normalization satisfying $b_n^2 \sim nH(b_n)$. In our case, $b_n^2 \sim 4n \ln b_n$, implying that $b_n^2 \sim 2n \ln n$.

For each n, let m_n be such that $T_{m_n} \le n < T_{m_n+1}$.

We have the following representation

$$\sum_{k=1}^{n} X_k - \sum_{k=1}^{[n/2]} Y_k = (T_0 - 1)X_0 + (\sum_{k=1}^{m_n} \tau_k X_{\tau_k} - \sum_{k=1}^{[n/2]} \tau_k X_{\tau_k}) + \sum_{k=T_{m_n+1}}^{n} X_k, \tag{31}$$

where $Y_k = \tau_k X_{\tau_k}$ is a centered i.i.d. sequence in the domain of attraction of a normal law. By the limit theorem for i.i.d. variables in the domain of attraction of a stable law [see Feller (1971)] we obtain,

$$\frac{\sum_{k=1}^{[n/2]} Y_k}{b_{[n/2]}} \Rightarrow N(0,1). \tag{32}$$

By Theorem 4.1 from Billingsley (1968) the CLT for $(\sum_{k=1}^{n} X_k)/b_{[n/2]}$ will follow from (31) and (32) provided we show that the normalized quantity in the right-hand side of (31) converges in probability to 0. Clearly, because $b_{[n/2]} \to \infty$ we have

$$\frac{(T_0 - 1)X_0}{b_{[n/2]}} \Rightarrow 0.$$

Also,

$$\mathbb{E}\frac{|\sum_{k=T_{m_n+1}}^{n} X_k|}{b_{[n/2]}} \le \frac{\mathbb{E}|\tau_{m_n+1}|}{b_{[n/2]}} = \frac{2}{b_{[n/2]}} \to 0.$$

Therefore it remains to study the middle term. Let $\delta > 0$.

$$\mathbb{P}(|\sum_{k=1}^{m_n} Y_k - \sum_{k=1}^{[n/2]} Y_k| > \varepsilon b_{[n/2]}) \le \mathbb{P}(|\frac{m_n}{n} - \frac{1}{2}| \ge \delta)$$

$$+\mathbb{P}(\max_{n/2-\delta n < l < n/2+\delta n} |\sum_{k=1}^{l} Y_k - \sum_{k=1}^{[n/2]} Y_k| > \varepsilon b_{[n/2]}) = I + II$$

Then, by the definition of m_n and the law of large numbers for the i.i.d. sequence $(\tau_i)_{i \ge 1}$ we know that: $m_n/n \to 1/\mathbb{E}(\tau_1) = 1/2$ a.s. Therefore the first term converges to 0 for every δ fixed as $n \to \infty$. As for the second term, by stationarity and the fact that Y_k are i.i.d.

$$II \le 2\ \mathbb{P}(\max_{1 \le l \le [\delta n]+1} |\sum_{k=1}^{l} Y_k| > \varepsilon b_{[n/2]}/2)$$

and Theorem 1.1.5 in De la Peña and Giné (1999),

$$II \leq 2\ \mathbb{P}(\max_{1\leq l\leq[\delta n]+1} |\sum_{k=1}^{l} Y_k| > \varepsilon b_{[n/2]}/2) \leq 18\ \mathbb{P}(|\sum_{k=1}^{[\delta n]+1} Y_k| > \varepsilon b_{[n/2]}/60).$$

Then, by the central limit theorem in (32) and the fact that $b_n^2 \sim 2n\ln n$, we have

$$\limsup_{n\to\infty} \mathbb{P}(|\sum_{k=1}^{[\delta n]+1} Y_k| > \varepsilon b_{[n/2]}/60) = \limsup_{n\to\infty} \mathbb{P}(|\sum_{k=1}^{[\delta n]+1} Y_k|/b_{[\delta n]} > \varepsilon b_{[n/2]}/60 b_{[\delta n]})$$
$$\leq \mathbb{P}(N(0,1) > \varepsilon\delta^{-1/2}/120)$$

which converges to 0 as $\delta \to 0$.

It follows that

$$\frac{S_n}{b_{[n/2]}} \Rightarrow N(0,1).$$

We recall that $\sigma_n^2 = 2n\log n = b_n^2$, implying that

$$\frac{S_n}{\sigma_n} \Rightarrow N(0,\frac{1}{2}).$$

Consequently, because the chain is irreducible and aperiodic, by Corollary 3, $W_n(t) \Rightarrow 2^{-1/2}W(t)$. $\Diamond$

For a different example having this type of asymptotic behavior we cite Zhao et al. (2010). Our Corollary 3 will also provide a functional central limit theorem for their example.

Acknowledgement. The authors would like to thank the referee for carefully reading the manuscript and for numerous suggestions that improved the presentation of this paper. The last author would like to thank Mikhail Gordin for suggesting the use of backward forward martingale decomposition.

References

[1] Billingsley, P. (1968). *Convergence of probability measures.* John Wiley, New York.

[2] Bradley, R.C. (2007). *Introduction to Strong Mixing Conditions.* vol. 1-3, Kendrick Press.

[3] Cuny, C. and Peligrad, M. (2012). Central limit theorem started at a point for additive functional of reversible Markov Chains. *J. Theoret. Probab.* **25**, 171-188.

[4] De La Peña, V. and Giné, E. (1999). *Decoupling: From dependence to independence: Randomly stopped processes, U-statistics and Processes, Martingales and Beyond.* Springer-Verlag, New York.

[5] Derriennic, Y. and Lin, M. (2001). The central limit thorem for Markov chains with normal transition operators started at a point, *Probab. Theory Relat. Fields*, **119**, 508-528.

[6] Doob, J. (1953). *Stochastic Processes.* Wiley, New York.

[7] Doukhan, P., Massart, P. and Rio, E. (1994). The functional central limit theorem for strongly mixing processes. *Ann. Inst. H. Poincaré Probab. Stat.* **30,** 63-82.

[8] Feller, W. (1971). *An Introduction to Probability Theory and its Applications.* Vol. II. John Wiley, New York.

[9] Kipnis, C. and Landim, C. (1999). *Scaling Limits of Interacting Particle Systems.* Springer, New York.

[10] Kipnis, C. and Varadhan, S.R.S. (1986). Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.* **104**, 1-19.

[11] Lyons, T.J. and Zheng, W.A. (1988). A crossing estimate for the canonical process on a Dirichlet space and a tightness result. Colloque Paul Lévy sur les processes stochastiques, *Asterisque* **157–158** 249–272.

[12] Merlevède, F. and Peligrad, M. (2011). Rosenthal-type inequalities for the maximum of partial sums of stationary processes and examples; to appear in *Ann. Probab.* DOI:10.1214/11-AOP694.

[13] Meyer, P.A. and Zheng, W.A. (1984). *Construction du processus de Nelson reversible.* Sém. Probab. XIX, Lect. Notes in Math. No 1123, 12–26.

[14] Rio, E. (2000). *Théorie asymptotique des processus aléatoires faiblement dépendants.* Mathématiques et Applications. **31,** Springer, Berlin.

[15] Rio, E. (2009). Moment inequalities for sums of dependent random variables under projective conditions. *J. Theoret. Probab.* **22,** 146-163.

[16] Tierney, L. (1994). Markov chains for exploring posterior distribution (with discussion). *Ann. Statist.* **22**, 1701-1762.

[17] Wu, L. (1999). Forward-backward martingale decomposition and compactness results for additive functionals of stationary ergodic Markov processes. *Ann. Inst. H. Poincaré Probab. Stat.* **35**, 121-141.

[18] Wu, W.B. and Woodroofe, M. (2004). Martingale approximations for sums of stationary processes. *Ann. Probab.* 32, 1674-1690.

[19] Zhao, O. and Woodroofe, M. (2008). Law of the iterated logarithm for stationary processes. *Ann. Probab.* **36**, 127-142.

[20] Zhao, O., Woodroofe, M. and Volný, D. (2010). A central limit theorem for reversible processes with nonlinear growth of variance, *J. Appl. Prob.* 47, 1195-1202.